DOCUMENT RESUME

ED 190 671 TM 800 479

AUTHOR Martin, Jeanne: And Others
TITLE Within-Class Analyses of Relationships between Student Achievement and Teacher Behaviors.
INSTITUTION Texas Univ., Austin. Research and Development Center for Teacher Education.
SPONS AGENCY National Inst. of Education (DHEW), Washington, D.C.
REPORT NO R&D-R-4075
PUB DATE Apr 80
CONTRACT NIE-C-741-0089
GRANT OB-NIE-G-78-0216
NOTE 31p.: Paper presented at the Annual Meeting of the American Educational Research Association (64th, Boston, MA, April 7-11, 1980).

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Academic Achievement: *Classroom Environment: Classroom Observation Techniques: Elementary School Teachers: Grade 1: Grouping (Instructional Purposes): Homogeneous Grouping: Primary Education: *Reading Readiness: *Student Teacher Relationship: *Teacher Behavior: *Teacher Effectiveness
IDENTIFIERS *First Grade Reading Group Study: Metropolitan Achievement tests: Metropolitan Readiness Tests

ABSTRACT

Data were collected in a teacher effectiveness study on the interaction of teachers and first grade students while the students participated in reading groups within intact classrooms. The teacher behavior variables were then related to students achievement using three units of analysis: class means, reading group means within classes, and student scores within reading groups. Results revealed that relationships differed depending upon the level of the analysis. Particularly striking was the large number of significant relationships which occurred only at the student-within-reading-group level. Apparently, some variables took on different meanings in the correct context of a small group, implying that students were very much aware of and affected by their learning environment. The results demonstrate that a combination of within-class and between-class analyses provides a better understanding of the relationship of teacher-student interactions to student achievement than either within-class or between-class analyses alone. (Author)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

Within-class Analyses of Relationships Between Student Achievement and Teacher Behaviors

Jeanne Martin and Donald J. Veldman

The Research and Development Center
for Teacher Education

The University of Texas at Austin

Linda M. Anderson

Michigan State University

R&D Rep. No. 4075

This study was supported in part by the National Institute of Education under Contract NIE-C-74-0089 and Grant OB-NIE-G-78-0216, The Research and Development Center for Teacher Education. The opinions expressed herein do not necessarily reflect the position or policy of the National Institute of Education and no official endorsement by that office should be inferred.

APR 2 3 1980

Abstract

Data were collected in a teacher effectiveness study on the interaction of teachers and first-grade students while the students participated in reading groups within intact classrooms. The teacher behavior variables were then related to student achievement using three units of analysis: class means, reading group means within classes, and student scores within reading groups. Results revealed that relationships differed depending upon the level of the analysis. Particularly striking was the large number of significant relationships which occurred only at the student-within-reading-group level. Apparently, some variables took on different meanings in the context of a small group, implying that students were very much aware of and affected by their learning environment. These results demonstrate that a combination of within-class and between-class analyses provides a better understanding of the relationship of teacher-student interactions to student achievement than either within-class or between-class analyses alone.

Within-class Analyses of Relationships Between Student Achievement and Teacher Behaviors

Several recent studies of teaching effectiveness have yielded results with implications for practice (e. g., see Borich, 1977; Brophy, in press; Brophy & Evertson, 1976; Evertson, Anderson, Anderson, & Brophy, in press; Good, 1979; Rosenshine, 1976). Some of the critical teacher behaviors identified through this research include interactions between the teacher and individual students within the class. For example, the frequency with which the teacher contacted individuals, the nature of the interactions (whether academic, behavioral, or social), the difficulty level of questions, and the type of feedback given to students' answers have all yielded significant correlations with student achievement. However, in most studies, data have been analyzed using class means, even though the data were recorded originally about the teacher's interactions with individual students.

Use of class means avoids the methodological problem of nonindependence which arises when data are analyzed with student scores pooled without regard to class membership. Recently, however, the use of class means to analyze data collected on individual students has received both methodological and substantive criticism. Cronbach (Note 1) argues that class level analyses can be misleading, because relationships obtained at the class level may not occur within classes. Likewise, one cannot assume that the absence of a relationship at the class level implies that there is no relationship within classes. Brophy (Note 2) pointed out that most teacher behaviors are directed to individual students rather than to the whole class. Even when targeted to the whole group, individual

differences among students may influence the outcome of the teacher's behavior. Therefore, Brophy argued that the student is the most meaningful unit of analysis.

Those arguments do not imply that analyses of class means have been useless. Through such analyses, a body of knowledge has developed about general patterns of teacher behaviors that are related to student outcomes. Such information could only be obtained by looking at a large sample of teachers and comparing their behaviors and effects, and this kind of examination required the use of class mean scores. However, it is also important to know about systematic patterns of within-class relationships between teacher behaviors and student learning.

One area of educational research that is dependent on the elucidation of within-class relationships is research on teacher expectations (reviewed by Braun, 1976, and by Brophy and Good, 1974). Expectation researchers have suggested that the teachers' different treatment of various students within a class is based, to a large extent, on the teachers' expectations for those students' academic success. They also suggest that such different treatment of students within classes may lead to student outcomes in line with the teachers' expectations. For example, if a teacher expects that a student will not be able to answer difficult questions, the teacher may ask the student fewer challenging questions and may push him less hard than other students in the class to practice important skills. Consequently, the student may achieve less than his peers because of fewer opportunities and less encouragement. In order to explore hypotheses such as these, researchers must be able to examine relationships within classes.

Other research that suggests important within-class influences on the relationship of teacher behaviors to student achievement is that of Weinstein and Middlestadt (1979). They found that students had fairly consistent perceptions of high and low achievers, and associated differing teacher behaviors with the two groups. This suggests that students' perceptions of their peers are related to the teacher's different treatment of high and low achievers. Weinstein and Middlestadt suggest that because classrooms are social settings, other students may contribute a great deal to an individual's perception of himself or herself as a learner, which in turn will influence effort and achievement.

Given such potential within-class influences on the interactions and achievement of individual students, analyses of within-class relationships (utilizing data at the student level) would be a useful supplement to between-class analyses that rely on class average scores.

Unfortunately, most large studies of teacher effectiveness that included achievement data can only be examined at the class level. However, the present paper reports further analysis of a large data base, the First-grade Reading Group Study (Anderson, Evertson, & Brophy, 1979), in which data were collected on the teachers' interactions with individual students while those students participated in reading groups within intact classrooms. Therefore, in addition to the conventional class mean analyses, two types of within-class analyses were performed with these data: one that considered small groups within classes as instructional units, and one that examined individual students' interactions within the small group. This paper discusses the substantive implications of the results of these hierarchical analyses.

Method

Overview

The First-grade Reading Group Study was an experimental study that tested the effectiveness of a model containing 22 principles of small group instruction. Ten treatment teachers agreed to apply the model when teaching their reading groups. Ten control teachers were not given the instructional principles. All 20 teachers were observed teaching reading groups throughout the year.

The Metropolitan Readiness Test was given at the beginning of the school year. A Total Readiness score was computed from six subscale scores and was used in all analyses as the pretest measure. The reading subtests of the Metropolitan Achievement Test, Level I, were given at the end of the year. Two subtests were used to create a Total Reading score, which was the posttest measure of student achievement in all analyses reported here.

Classroom observations were recorded with a low-inference coding system developed specifically for the study. Only data describing academic interactions with individual students initiated by the teacher during reading group instruction are reported in this paper.

The methodology and results of the study are completely described in Anderson, Evertson, and Brophy (1979, Note 3).

Subjects

Data were collected on 20 teachers in six middle-SES schools serving a predominantly white population in an urban district in the Southwest. The original sample of 20 teachers was reduced in order to obtain a consistent set of data for the analyses presented in this paper. It was necessary to have a data set that contained observation scores recorded for students while they participated in a single reading group within a single intact

classroom throughout the year. However, there was student mobility between groups, some students had incomplete data, and some teachers formed new groups that existed for only a few observations. Hence, the results presented in this investigation are based on a subset of the original sample. Students with incomplete test data were not considered. Reading groups that were observed less than six times or contained less than four students were eliminated from analysis.

Most teachers indicated that they formed their reading groups on the basis of the Metropolitan Reading Readiness scores, so that students within reading groups were relatively homogenous with respect to ability level. However, because teachers formed reading groups on the basis of ability, the reading groups within a class were heterogenous with respect to ability level--there was generally a high, a middle, and a low ability group in each class. Classes which contained only low or high reading groups (due to team teaching) or classes with less than two reading groups were dropped from the study. The end result was a set of data in which every student participated in a single reading group within a single classroom for the five months of the observation.

The original data set consisted of 20 classes, 67 reading groups, and 540 students. This final data set consisted of 15 classes, 41 reading groups, and 288 students. Of the 15 classes, nine were treatment classes and six were control classes. The number of reading groups in each class varied from two to four with a mean of 2.7 groups per class. The number of students in each reading group varied from four to 13 with a mean of seven students per group.

Data Collection

Four observers were trained to use the observational coding system.

Every observer coded with every other observer until 80% agreement was reached on each major section of the coding system. After this time, observers worked alone. Each of the 20 teachers was seen by two observers who alternated visits to classrooms, beginning in November and continuing about once a week through April, so that each class was observed 15 to 20 times during the year.

Independent Variables

The variables reported here were chosen because they represented the four major components of an academic interaction between the teacher and an individual student: selection of student, questioning by the teacher, the students' answer, and the teacher's response to the answer.

Proportion variables were created from several categories of the observation system. Frequencies of single categories (e.g., the number of correct answers) were summed across all observations for each unit of analysis (teachers, groups, and students). Then, these frequencies were used to create proportions representing actual occurrences compared to maximum possible occurrences (e.g., the proportion of all answers that were correct answers). These proportion variables are listed in Table 2.

Data Analyses

For all three levels of analyses, tests were performed to detect statistical interactions (i.e., the relationships between teacher behavior and achievement varied according to entering readiness) and main effects (i.e., there was a linear relationship between teacher behavior and achievement when entering readiness was a covariable).

Each of the three levels of analyses addressed different questions, and significant main effects should be interpreted in these ways:

1) When class means were involved, relationships described ways that teacher behaviors, which differed between classes, related to differences in class-mean achievement. 2) In contrast, reading-group-within-class analyses determined whether reading group means on teacher behaviors were related to group mean adjusted achievement. Because reading group means were compared within single classes, relationships obtained in these analyses were relative. That is, a positive relationship indicated that reading groups that received a higher amount of the teacher behavior relative to other groups within the same class achieved more--regardless of the actual amount of the teacher behavior received by the higher groups. No conclusions can be drawn about the "optimal" amount of teacher behavior. 3) Similarly, individual-within-reading-group analyses revealed whether student scores were related to adjusted student achievement. Because student scores were compared within reading groups, relationships obtained in these analyses were also relative. That is, a positive relationship indicated that students who received more of that teacher behavior, relative to other students in their reading groups, achieved more. These results do not indicate what level of teacher behavior was associated with higher achievement.

Analyses involved three sets of regression models that were evaluated for each independent variable. In all models, reading readiness scores were used as a covariable and reading achievement scores were used as the criterion. (For more information on the development of the regression models, see Martin, Note 4.)

The first set of models (shown below in abbreviated notation) contained class means as the unit of analysis. Predictors were the teacher

behavior variable and the interaction of the teacher behavior variable with readiness:

$$Y_c = R_c + TB_c + TB_c*R_c + E_c,$$

where Y_c is the reading achievement score, class mean,

R_c is the reading readiness score, class mean,

TB_c is the teacher behavior described by that variable, class mean,

TB_c*R_c is the vector describing the interaction of readiness and the teacher behavior, and

E_c is the error term.

The second set of models contained reading group means as the unit of analysis. Predictors were the teacher behavior variable, the interaction of the teacher behavior variable with readiness, and a set of binary vectors indicating class membership. Inclusion of the binary vectors in the restricted models in which TB_g*R_g and then TB_g were omitted has the effect of analyzing reading group scores within classes.

$$Y_g = R_g + TB_g + TB_g*R_g + C^1 + C^2 + \ldots + C^{15} + E_g$$

where C_1, C_2, etc. are binary class membership vectors. Other terms correspond to the definitions given above, except that all entries were group means.

The third set of models contained student scores as the unit of analysis. Predictors were the teacher behavior variable, the interaction of the teacher behavior variable with readiness, and a set of binary vectors denoting reading group membership. Inclusion of the binary vectors allowed student scores to be analyzed within reading groups.

$$Y_s = R_s + TB_s + TB_s*R_s + G^1 + G^2 + \ldots + G^{41} + E_s$$

where G_1, G_2, etc. are binary group membership vectors. Other terms correspond to the definitions given above, except that all entries were student scores.

Results

The average means and standard deviations for reading groups within classes and students within reading groups for both readiness and achievement scores are illustrated in Table 1. Students within reading groups were more homogenous with respect to both achievement and readiness than reading groups within classes. This lends confirmation to the teachers' assertions that they formed reading groups primarily on the basis of readiness scores and also provides information necessary for interpreting the results presented below.

Significant main effects obtained from the regression equations for all three levels of analysis are presented in Table 2. Significant interactions are described in the footnotes below the table. Results are summarized below for each level of analysis. It should be noted that the results may be affected by the fact that nine of the 15 classes in this study are treatment classes. These nine teachers were given a treatment manual which stressed the importance of certain teacher behaviors and management principles. No effort was made in the treatment manual to encourage teachers to treat reading groups of different ability levels differently, i.e., the principles presented in the manual were to be used with all reading groups. Therefore, the variance for certain teacher behaviors may be artificially reduced at the reading-group-within-class and the student-within-reading-group levels of analysis in the nine treatment classes. However, the variance at the class level may be increased due to the increase in usage of certain teacher behaviors by the treatment

classes. This potential change in variance due to the inclusion of treatment classes should not affect the relationship of teacher behaviors to achievement, but may make detection of within-class relationships more difficult.

Results of Class-level Analyses

There were four significant findings when class means were used. The following were related to higher class mean achievement: higher levels of correct answers (Variable 10), lower levels of failures to respond (Variable 13), greater use of new (typically related) questions following correct answers (Variable 17), and less use of giving the answer when a student failed to respond (Variable 23).

These results suggest that classes in which more students had higher proportions of successful interactions, and in which the teacher showed a greater tendency to sustain interactions, had greater achievement. This interpretation is supported by other significant findings at the class level reported in Anderson et al. (1979, Note 3).

Results of Group-level Analyses

These analyses examined differences between groups within classes to determine if differential teacher treatment of the groups was related to different adjusted achievement. Three of the significant results matched those at the class level: Higher achievement was related to higher proportions of correct answers (Variable 10), fewer failures to respond (Variable 13), and lower rates of giving the answer to a student after he failed to respond (Variable 23). One other simple effect emerged: Groups in which the teacher selected students more often by calling on nonvolunteers in a random fashion had lower achievement (Variable 5). This variable may represent something about the eagerness of group members to

participate, or the attention level of students in the group. Teachers often use this selection technique when students do not volunteer, when there is no consistent pattern for calling on students equally, or when they need to bring a student back to attention. If these conditions existed in one group more than another within a class, and if they led to higher rates of nonvolunteer selection, then it is reasonable that the variable would be associated with lower achievement.

There was one significant interaction with entering ability for the use of a clue (usually through simpler questions) following incorrect answers (Variable 22). Within a class, reading groups with higher achievement tended to perform better when provided with a clue following an incorrect response than reading groups with lower achievement.

In summary, the group analyses confirm some of the same patterns seen at the class level regarding students' successful interactions and the importance of the teacher's response to a student who has failed to respond. For these variables, the within-class relationships, at least at the level of the instructional group, are the same as the class-level relationships.

Another important characteristic of the data concerns proportions of variance at the different levels of analysis. An examination of the average variance for teacher behaviors reveals considerably less variance at the reading-group-within-class level than at the class level for the following variables: teacher selection variables, failure to provide feedback to correct responses, providing emphasis feedback to correct answers, and for providing terminating and sustaining feedback to incorrect answers. The small variance at the reading group level indicates that teachers tended to provide similar proportions of these behaviors to all of

the reading groups in their class regardless of the ability level of the groups. This may be due to the effect of the treatment (as discussed previously), or it may reflect a real tendency by teachers to respond similarly to all reading groups regardless of ability level, at least with respect to these variables.

Results of Student-level Analyses

These results reflect patterns present across all reading groups. Significant results suggest that the teacher's different treatment of students within a group is related to the different achievement of those students.

Fifteen of the 26 variables yielded significant results. However, the actual differences in R^2 were not large. Generally, only 1% or 2% additional variance was predicted by the teacher behavior variable at this level. Therefore, results should be interpreted cautiously, even though the data reveal meaningful patterns.

Three of the relationships significant at the student-within-group level of analysis were also significant at the reading-group-within-class and class levels of analysis. These relationships indicated that students' achieved more when they answered more questions correctly, had fewer failures to respond, and had fewer response opportunities terminated by the teacher giving the student the answer. Because the relationships were consistent across all levels of analysis, they indicate that within the range of response observed in these data, the greater the amount of these behaviors, the better (or worse) the students achieved.

The rest of the relationships presented in this section are those that were obtained only at the student-within-reading-group level of analysis. The presence of a relationship only at this level suggests that

within-group comparisons were responsible for the significant relationship. In other words, the absolute amount of the behavior a student received was not related to achievement (which may be why the relationship did not show up at the class or reading group levels of analysis), but the relative standing of the student within his group was related to achievement. That is, it did not matter how much of the behavior a student received, but rather how much he received relative to his peers in the same group.

The rate of academic interactions (per minute of lesson time) was negatively related to adjusted achievement within the group. Perhaps, within a group, students who were called on most were called on to maintain their attention, and, if so, inattention may have been at the root of their lower achievement.

Analyses of variables describing the nature of teacher-student interactions suggest that quality of contacts is also an important predictor of achievement, when examined on the student level. For these variables, each student's frequencies were divided by the total number of interactions with the teacher, so there was a control for differences among students in frequency of interactions. Results are summarized below:

1. Students who had higher achievement, relative to others within their reading groups, were less likely to have called out the answer without permission (Variable 7). At the reading group and class levels of analysis, fewer call outs were not associated with higher achievement. This suggests that the relationship at the student level may be due to forces occurring within the group. For example, the student who calls out answers may be affected adversely by the response of the teacher or other students in the group to call outs. Accordingly, call outs themselves may not be detrimental to achievement (as evidenced by the lack of a

relationship at the reading group and class levels of analysis), but when they occur in the context of a small group, other phenomena may occur that cause call outs to relate negatively to achievement.

2. Students who achieved more were asked more nonreading questions, relative to their lower-achieving peers (Variable 9). Nonreading questions included comprehension questions, "thought" questions, or fact questions, and excluded questions that required students to decode words and focus on sounds. A simple explanation, such as nonreading questions provide more practice on tested skills, is inapplicable because of the absence of a relationship at the reading group and class levels of analysis. Rather, an explanation that considers the dynamics of the small group is more appropriate. Cooper (1979) reviewed several studies that indicated teachers try to establish a pattern of successful interactions and consequently, may direct more difficult questions to higher achieving students and less difficult questions to lower achieving students. Perhaps, within reading groups, the teacher directs the more involved nonreading questions to the higher achieving students and the simpler reading questions to the lower achieving students. If the students in the group are aware of this teacher differentiation, lower achieving students may respond by developing more negative feelings about their academic behavior and may work less hard.

3. Students who achieved more had higher proportions of correct answers (Variable 10) and lower levels of incorrect answers (Variable 11), "don't know" answers (Variable 12), and failures to respond (Variable 13). That is, more of their interactions were successful.

4. The feedback given to students' answers by the teacher yielded several relationships with relative achievement.

a) Praise after correct answers interacted with entering ability, even within ability-based groups (Variable 16). Students with higher entering readiness did worse with higher levels of praise, whereas the opposite was true of students with lower entering readiness. This finding is in keeping with other results at the class level that suggested that lower ability students do better with more encouragement (Brophy & Evertson, 1976). However, it is surprising to find similar interactions within groups that are supposedly homogeneous.

b) Students achieved less when they received more feedback to incorrect answers in the form of criticism (Variable 18) or "terminal" feedback (Variable 19). Terminal feedback was coded whenever the teacher reacted to an error or a failure to respond by either giving the students the answer or by having another student supply the answer.

c) Students who were more often able to improve their incorrect answers through "sustaining" feedback achieved more than their peers (Variable 20). Sustaining feedback usually meant repeating the question and allowing more time or giving clues to simplify the question.

In summary, these results suggest that the students within the group who had more successful interactions, who were asked more nonreading (e.g., comprehension) questions, and who were more often given chances to improve incorrect answers, achieved more. In effect, these students had more positive learning experiences and such experiences may have provided students with greater confidence and encouragement. In contrast, students who achieved less, relative to others in their groups, had less successful interactions, displayed a greater tendency to shout out answers before being called on, and received more criticism. Overall, these students appear to have had more negative learning experiences and such experiences

may have provided these students with less encouragement. Cooper (1979) provides some support for this notion in his review of expectation research; he suggests that less successful students may develop feelings of less efficacy and consequently, may exhibit less motivation.

An examination of the variances for the teacher behavior variables at the student-within-reading-group level reveals that teachers tended to treat all students within a reading group similarly with respect to the following variables: teacher selection variables, emphasis feedback to correct answers, and lack of feedback to correct answers. These variables also had restricted variance at the reading-group-within-class level, suggesting that these variables represent behaviors that are characteristic of a teacher's general instructional style and therefore, remain relatively stable within classes. Again, this may merely reflect an artifact of the treatment classes and may not be representative of actual classrooms.

There were also a number of variables which showed greater variance at the student-within-reading-group level of analysis than at either the class or reading-group-within-class levels. These variables included proportion of correct answers and various types of sustaining and terminating feedback. Apparently, these particular behaviors do not represent a characteristic teacher style but instead are triggered by the teachers' perception of individual students or, in the case of correct answers, may result from individual student differences within the group.

Conclusions

Several variables revealed similar relationships across all three levels of analyses, suggesting that the more or less of these teacher behaviors a student received, the more the student achieved. One cluster of variables that was significant at each level suggested that a pattern of

instruction including more overall contacts that lead to successful practice of reading skills is likely to maximize achievement. This conclusion is not new; many other investigations of class-level data have also suggested this. However, the discovery of significant relationships within classes highlights the importance of these variables.

Significant findings only at the student-within-group level suggest that the context in which teacher behaviors occur is an important determinant of their effect upon achievement. Within a group, it appears that the relative standing of a student with respect to his peers on certain behaviors may be more important than the actual amount of the behavior that the student receives. It was suggested that this may be due to the students' awareness of his standing within the group. Those students who achieve less than their peers, regardless of the overall ability level of the group, generally have more negative academic interactions and those interactions may lead to feelings of inadequacy and decreased motivation. Conversely, those students who achieve more than their peers within the same group generally have more positive academic interactions and these interactions may lead to greater confidence and achievement motivation.

Although the variables in this study have been discussed in terms of teacher behaviors, thereby implying they are under the teacher's control, it is important to point out that they may also refer to student differences that could influence student achievement. For example, student level of correct answers, although certainly a function of the teacher's choice of questions, might also reflect student attention to the lesson and effort at applying the information. Students' calling out answers has already been discussed as a indication of impulsiveness. The higher number

of questions directed to students who eventually achieved less might indicate greater need for the teacher to monitor and control attention. Such variables, although termed "teacher behaviors," also reflect student characteristics that directly contribute to student achievement, and that may also influence teacher behavior.

There are several contextual factors which must be considered when interpreting results from this study. First, different results might be expected if the data were analyzed within whole class rather than small group settings. Some variables describing teacher contacts with students may assume different meanings depending upon the setting in which they occur. For example, teacher contacts with students may serve as a device to control attention in whole class settings but not in small groups where it is generally easier to get and hold student attention.

Second, the age level of the students may affect the results. Personal contacts with the teacher may be much more important in terms of student learning for first-graders who have not yet developed good listening skills than for junior high or high school students. Older students are probably much more capable of learning from hearing other students interact with the teacher (Brophy & Evertson, 1976; Evertson et al., in press).

Third, the effect of more general teacher characteristics (i.e., class level measures) on within-class relationships must be considered. It seems reasonable that the meaning of important behaviors (e.g., praise or criticism) might be different in classes where the behavior occurred frequently than in classes where it rarely occurred. In other words, research needs to be done to determine the importance of the absolute

amount of teacher behaviors on within-class relationships as well as the relative amounts of teacher behaviors.

Finally, the reader is again cautioned that this is an experimental study, not naturalistic. As mentioned previously, this could affect the interpretation of the results.

Obviously, many factors must be considered before a complete understanding of the effects of teachers' behaviors on student achievement will be accomplished. The analyses reported here have illustrated the importance of considering both within-class and within-group relationships in addition to the more traditional class level relationships.

Reference Notes

1. Cronbach, L. Research on classrooms and schools: Formulation of questions, design, and analysis. Occasional paper of the Stanford Evaluation Consortium, Stanford University, 1976.
2. Brophy, J. The student as unit of analysis. Paper presented at the Annual Meeting of the American Educational Research Association, New York, April, 1977.
3. Anderson, L., Evertson, C., & Brophy, J. The First-grade Reading Group Study: Technical report of experimental effects and process-outcome relationships (R&D Rep. No. 4070). Austin, Texas: Research and Development Center for Teacher Education, The University of Texas at Austin, 1979.
4. Martin, J. The unit of analysis problem in teacher effectiveness research. Unpublished doctoral dissertation, University of Texas at Austin, 1978.

References

Anderson, L., Evertson, C., & Brophy, J. An experimental study of effective teaching in first-grade reading groups. Elementary School Journal, 1979, 79, 193-223.

Borich, G. (Ed.). The appraisal of teaching: Concepts and process. Reading, Massachusetts: Addison-Wesley, 1977.

Braun, C. Teacher expectation: Sociopsychological dynamics. Review of Educational Research, 1976, 46, 185-213.

Brophy, J. Teacher behavior and its effects. Journal of Educational Psychology, in press.

Brophy, J., & Evertson, C. Learning from teaching: A developmental perspective. Boston: Allyn and Bacon, Inc., 1976.

Brophy, J. & Good, T. Teacher-student relationships: Causes and consequences. New York: Holt, Rinehart, & Winston, 1974.

Cooper, H. M. Pygmalion grows up: A model for teacher expectation communication and performance influence. Review of Educational Research, 1979, 49, 389-410.

Evertson, C., Anderson, C., Anderson, L., & Brophy, J. Relationships between classroom behaviors and student outcomes in junior high mathematics and English classes. American Educational Research Journal, in press.

Good, T. Teacher effectiveness in the elementary school. Journal of Teacher Education, 1979, 30, 52-64.

Rosenshine, B. Classroom instruction. In Gage, N. (Ed.), The psychology of teaching methods, Seventy-fifth yearbook of the National Society for

the Study of Education (Part I). Chicago: University of Chicago Press, 1976.

Rosenshine, B. V., & Berliner, D. C. Academic engaged time. British Journal of Teacher Education, 1978, 4, 3-16.

Weinstein, R., & Middlestadt, S. E. Student perceptions of teacher inceractions with male high and low achievers. Journal of Educational Psychology, 1979, 71, 421-431.

Acknowledgements

The authors thank Carolyn Evertson, Director of the Correlates of Effective Teaching Program at The Research and Development Center for Teacher Education, The University of Texas at Austin, without whose support this investigation would not have been possible.

Table 1

Average Means and Standard Deviations

for the Three Levels of Analysis

	Mean[1]		Standard Deviation[2]	
	Readiness	Achievement	Readiness	Achievement
Class	68.03	57.27	10.75	9.32
Reading Group Within Class	67.89	57.10	9.21	10.38
Student Within Reading Group	66.53	55.95	6.81	7.22

[1]Reading group means were averaged for each class and then these class means were averaged across classes. Similarly, student scores were averaged for each reading group and then these reading group means were averaged across reading groups and classes.

[2]Standard deviations of reading group means were computed within each class and then averaged across classes. Similarly, standard deviations of student scores were computed within each reading group and then averaged across reading groups and classes.

Table 2

Summary Table of Results Across All Levels of Analysis

Teacher Behavior	Class $\bar{x}$	Class SD	Class Difference in R^2	Reading Group[1] $\bar{x}$	Reading Group[1] SD	Reading Group[1] Difference in R^2	Student[2] $\bar{x}$	Student[2] SD	Student[2] Difference in R^2
1. Response opportunities per minute	.29	.07	.0297	.31	.12	.0004	.29	.13	.0217[a](-)
Teacher selection of students:									
2. Based on ordered turns	.54	.29	.0093	.53	.07	.0136	.53	.07	.0003
3. Prior to asking a question	.08	.10	.0269	.09	.02	.0025	.09	.03	.0000
4. Who are nonvolunteers	.18	.12	.0002	.18	.03	.0154(-)	.19	.06	.0000
5. Who are volunteers	.14	.11	.0028	.14	.03	.0065	.13	.05	.0018
6. Student call outs in response to a teacher question	.06	.04	.0028	.06	.02	.0010	.06	.05	.0070(-)
Response opportunities:									
7. That were reading questions	.69	.14	.0002	.68	.06	.0006	.69	.06	.0066(+)
8. That were non-reading questions	.30	.14	.0003	.30	.06	.0009	.30	.06	.0050(+)
9. With correct answers	.71	.06	.1054(+)	.72	.06	.0201(+)	.72	.10	.0737(+)
10. With incorrect answers	.22	.05	.0226	.21	.05	.0102	.21	.08	.0493(+)
11. With don't know answers	.02	.01	.0033	.01	.01	.0007	.01	.01	.0163(-)

Table 2-continued

Teacher Behavior	Class $\bar{x}$	Class SD	Class Difference in R^2	Reading Group[1] $\bar{x}$	Reading Group[1] SD	Reading Group[1] Difference in R^2	Student[2] $\bar{x}$	Student[2] SD	Student[2] Difference in R^2
12. With no response answers	.12	.04	.1383(-)	.12	.03	.0169(-)	.12	.07	.0353(-)
Correct answers with:									
13. No feedback	.08	.08	.0329	.08	.04	.0002	.07	.03	.0001
14. Emphasis feedback	.37	.12	.0576	.36	.07	.0000	.37	.08	.0033
15. Praise feedback	.10	.06	.0464	.10	.05	.0005	.10	.05	.0004
16. New question feedback	.19	.04	.1013(+)	.19	.06	.0009	.18	.06	.0009
Incorrect answers with:									
17. Criticism feedback	.02	.01	.0178	.02	.02	.0002	.02	.03	.0048(-)
18. Terminating feedback	.49	.14	.0332	.50	.06	.0001	.49	.15	.0147(-)
19. Sustaining feedback	.48	.13	.0237	.46	.06	.0017	.48	.14	.0081(+)
Incorrect answers with:									
20. Give answer feedback	.35	.14	.0450	.35	.14	.0097	.35	.22	.0079(-)
21. Clue feedback	.24	.11	.0114	.24	.12	.0004	.23	.18	.0017
22. Improvement	.34	.10	.0383	.34	.12	.0000	.35	.20	.0130(+)
No response answers with:									
23. Give answer feedback	.34	.16	.1368(-)	.33	.18	.0311(-)	.31	.27	.0187(-)

Table 2-continued

Teacher Behavior	Class $\bar{x}$	Class SD	Class Difference in R^2	Reading Group[1] $\bar{x}$	Reading Group[1] SD	Reading Group[1] Difference in R^2	Student[2] $\bar{x}$	Student[2] SD	Student[2] Difference in R^2
24. With clue feedback	.30	.13	.0257	.30	.15	.0000	.30	.24	.0003
25. Improvement	.33	.12	.0207	.33	.15	.0037	.36	.26	.0020

[1]For reading group results, means were calculated for each reading group within a class and then averaged across classes. Standard deviations of reading group means were also computed within each class and then averaged across classes.

[2]For individual student results, scores were calculated for each student and then averaged across reading groups and classes. Standard deviations of student scores were also computed within each reading group and then averaged across reading groups and classes.

[a]For each R^2 difference, probability values are indicated by underlining. $p < .05$ where one line appears; $p < .01$ where two lines appear.